

## Base Flipping in V(D)J Recombination: Insights into the Mechanism of Hairpin Formation, the 12/23 Rule, and the Coordination of Double-Strand Breaks<sup>∇†</sup>

Julien Bischerour,<sup>1‡</sup> Catherine Lu,<sup>2</sup> David B. Roth,<sup>2</sup> and Ronald Chalmers<sup>1\*</sup>

University of Nottingham, School of Biomedical Sciences, Medical School, Queen's Medical Centre, Nottingham NG7 2UH, United Kingdom,<sup>1</sup> and Kimmel Center for Biology and Medicine of the Skirball Institute and Department of Pathology, New York University School of Medicine, New York, New York 10016<sup>2</sup>

Received 10 February 2009/Returned for modification 3 March 2009/Accepted 21 August 2009

**Tn5 transposase cleaves the transposon end using a hairpin intermediate on the transposon end. This involves a flipped base that is stacked against a tryptophan residue in the protein. However, many other members of the cut-and-paste transposase family, including the RAG1 protein, produce a hairpin on the flanking DNA. We have investigated the reversed polarity of the reaction for RAG recombination. Although the RAG proteins appear to employ a base-flipping mechanism using aromatic residues, the putatively flipped base is not at the expected location and does not appear to stack against any of the said aromatic residues. We propose an alternative model in which a flipped base is accommodated in a nonspecific pocket or cleft within the recombinase. This is consistent with the location of the flipped base at position –1 in the coding flank, which can be occupied by purine or pyrimidine bases that would be difficult to stabilize using a single, highly specific, interaction. Finally, during this work we noticed that the putative base-flipping events on either side of the 12/23 recombination signal sequence paired complex are coupled to the nicking steps and serve to coordinate the double-strand breaks on either side of the complex.**

Antibody and T-cell receptor (TCR) diversity is generated by V(D)J recombination initiated by the RAG proteins, RAG1 and RAG2. The recombination signal sequences (RSSs), where recombination takes place, have a distinctive arrangement resembling transposon ends. The relationship between V(D)J recombination and transposition was established beyond doubt by the discovery of RAG-mediated transposition and by the identification of a triad of conserved active-site residues. This evidence placed RAG1 firmly within the family of transposases and retroviral integrases that have a characteristic DDE triad of amino acid residues that coordinate catalytic metal ions in the active site (1, 26, 30, 35, 39, 46). Later, the *Transib* family of transposons was identified as the likely ancestral group of RAG1 (33).

In V(D)J recombination, the RAG proteins excise the DNA between a pair of RSSs. This fragment is the equivalent of an excised transposon, and it takes no further part in the canonical V(D)J recombination reaction. Instead, the variable regions of the genes encoding antibodies and TCR are created by the imprecise rejoining of the flanking DNA, referred to as the “coding flank.” A key feature of the cleaved coding flanks is that they have covalently closed hairpin ends. The asymmetric resolution of these hairpins contributes to the diversification of the coding sequences during rejoining. The hairpins them-

selves arise as a consequence of the molecular mechanism RAG-mediated RSS cleavage.

The crystal structure for the catalytic core of the human immunodeficiency virus type 1 integrase protein revealed a structural fold shared in common with RNase H and the Holliday junction resolving enzyme RuvC (22). RNase H and RuvC monomers each perform a simple nicking reaction that requires a single phosphoryl transfer event. Cut-and-paste transposition, which requires at least three phosphoryl transfer steps at each transposon end, therefore presents a mechanistic challenge. One solution to this challenge was revealed by the discovery of the DNA-hairpin cleavage-intermediate in V(D)J recombination and *Tn10* transposition (Fig. 1) (34, 57). However, it is interesting to note that the existence of this intermediate was first suggested by Coen and colleagues on the basis of the genomic scars produced by excision of the hAT family transposon *Tam3* in *Antirrhinum majus* (14).

All DDE family transposases, including RAG1, cut the DNA to expose the 3'-OH at the end of the element (or RSS). However, the fate of the opposite strand and the order of strand cleavage events vary within the group (reviewed in references 13, 18, and 55). Some enzymes, such as the retroviral integrases and the bacteriophage Mu transposase, nick and integrate the 3'-OH directly without second-strand cleavage. The cut-and-paste transposons, which cleave both strands of DNA, can be divided into two groups. With some notable exceptions such as the *piggyBac* element, most prokaryotic family members cleave the bottom strand of the recombination site first, whereas most eukaryotic members cleave the top strand first (8, 10, 20, 41, 47, 48, 77). For those family members with a hairpin mechanism, the inverted polarity of the first step dictates the reversal of all subsequent steps (Fig. 1). In consequence, most eukaryotic members of the family can achieve

\* Corresponding author. Mailing address: University of Nottingham, School of Biomedical Sciences, Medical School, Queen's Medical Centre, Nottingham NG7 2UH, United Kingdom. Phone: 44 115 823 0087. Fax: 44 115 823 0103. E-mail: chalmers@nottingham.ac.uk.

† Supplemental material for this article may be found at <http://mcb.asm.org/>.

‡ Present address: Centre de Genetique Moleculaire, CNRS UPR2167, Avenue de la Terrasse, 91198 Gif sur Yvette Cedex, France.

<sup>∇</sup> Published ahead of print on 31 August 2009.

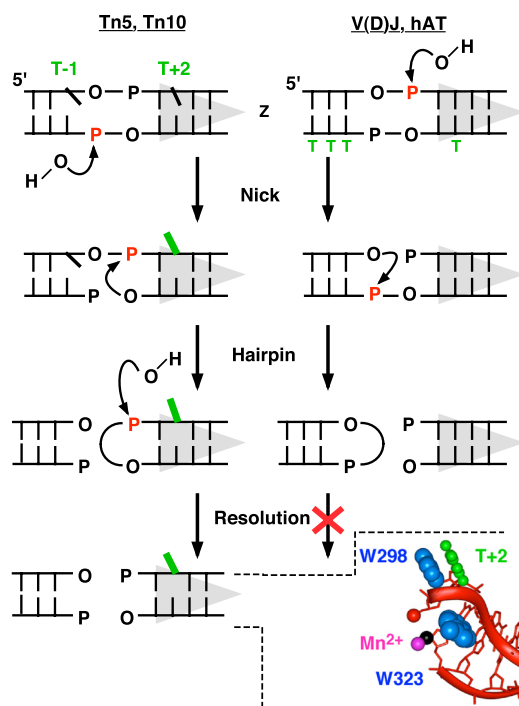


FIG. 1. Hairpin-processing reactions of opposite polarity. Most prokaryotic and eukaryotic members of the DDE family have hairpin intermediates of opposite polarity. In this paper, we refer to the two strands of DNA as “first strand” or “second strand” depending on the order of cleavage. The first strand therefore corresponds to the transferred and nontransferred strands of the prokaryotic and eukaryotic elements, respectively. Scissile phosphates are in red. The transposon end and RSS are shown as gray triangles. (Left panel) In Tn5 and Tn10, the first step of the reaction is a nick on the bottom (first) strand that exposes the 3'-OH at the end of the transposon. The second strand is cleaved by a direct transesterification reaction, which generates a “proximal-hairpin” intermediate on the transposon end (5, 34). Resolution by a nick at the tip of the hairpin yields a blunt transposon end. Distortion of the DNA helix can be detected by permanganate sensitivity of the T-1 and T+2 residues on the second strand. The insert shows the crystal structure of the Tn5 transposon end, highlighting the flipped base at position +2 (19). Two tryptophan residues are also shown. One acts as a “wedge” or “probe” residue inserted into the DNA helix, while the other provides stacking interactions that stabilize the flipped base. The W323 probe residue resides within the catalytic core close to the DDE residue E326, whereas the W298 stacking residue is in the inserted subdomain (see text for further details). Base flipping takes place after the first nick and is probably maintained for all subsequent steps, including integration (3, 7). (Right panel) In V(D)J recombination and the hAT family of transposons, the polarity of the reaction is reversed. The first nick is on the top strand providing a 3'-OH group on the flanking DNA (53, 71, 77). Transesterification yields a “distal hairpin” intermediate on the flanking DNA that is processed by the host. The positions of relevant thymidine residues in our substrates are indicated.

transposition with one less phosphoryl transfer reaction than the prokaryotic members, which are obliged to resolve the hairpin intermediate. The eukaryotic members can simply release the hairpin ends or, as in the case of RAG, hand them on to host factors for further processing (40).

Insight into the hairpin mechanism was provided by a crystal structure for the Tn5 transpososome, in which the penultimate base on the second, nontransferred, strand was flipped from the helix and stacked against a tryptophan side chain in the

protein (Fig. 1) (19). The flipped base seemed to provide the steric freedom that is presumed to be required for making and resolving the hairpin intermediate. Two groups searched for a residue in RAG1 that performs a function equivalent to the stacking tryptophan in the Tn5 transposase (27, 45). This work identified several candidate residues on the basis of their respective mechanistic defects and their rescue by modified DNA substrates.

Here we have further assessed the candidate stacking residues using biochemical techniques previously used to study the dynamics of base flipping in Tn5 and Tn10 transposition (6, 7). We have identified a distortion at position -1 of the V(D)J coding flank DNA. It is introduced after the first nick at the RSS and is therefore reminiscent of the flipped base at the end of Tn5. The distortion is perfectly correlated with the ability of wild-type and mutant RAG-RSS complexes to perform the hairpin step of the reaction. We conclude that this base is probably equivalent to the flipped base in Tn5. However, none of the candidate aromatic residues seems to fulfill the function of the putative stacking tryptophan residue. We therefore propose a model in which base flipping in RAG recombination is significantly different from that in Tn5 transposition.

Canonical V(D)J recombination occurs within a 12/23 RSS paired complex (24, 36, 60, 72, 73). This restriction is known as the 12/23 rule. More recently a further restriction, the so-called “beyond 12/23” (B12/23) rule has been proposed to explain the exclusion of direct V $\beta$ -to-J $\beta$  joining in the TCR  $\beta$  region, despite the presence of appropriately oriented pairs of 12 and 23 RSSs (4, 21, 31, 32).

Little is known of the mechanisms that enforce the 12/23 rule or coordinate cleavage on either side of the complex. However, during this work, we observed that the coding flank distortion was coupled on either side of a 12/23 RSS paired complex: the distortion of a nicked coding flank is suppressed by an unnicked partner. We present a model and discuss the biological significance of this conformational coupling and its relevance to the B12/23 rule.

## MATERIALS AND METHODS

**DNA substrates.** The oligonucleotides DAR39, DAR40, DAR42, DG61, DG62, DG4, and DG10 were used in this work, as described in reference 46. The following additional oligonucleotides were used (with sequences given in parentheses): QME27 (5'-TAAGACGTCGACGCT), QME28 (5'-[IdU]AAGACGTCGACGCTG), QME29 (5'-GGATCCGGTTTTTGTTCAGGGCTGTATCACTGTG), QME211 (5'-GGATCCGGTTTTTGTTCAGGGCTGTATCACTG[IdU]G) (25), WTTOP (5'-ACGCGTCGACGCTTACACAGTGATACAGCCTGAACAAAAACCGGATCC), TOPFLANK (5'-ACGCGTCGACGCTCTA), TOPARM (5'-CACAGTGATACGCCCTGAACAAAAACCGGATCC), OLIGO-LIG (5'-ACGCGTCGACGCTTACACAGTGATACAGCCCTGAACAAAAACC), DAR39AAA (5'-GATCTGGCCTGTCAAACACAGTGCTACAGACTGGAACAAAAACCTGCAG), DAR42AAA (5'-GATCTGGCCTGTCAAACAGACTGGAACAAAAACCTGCAG), DAR40TTT (5'-CTGCAGGGTTTTTGTTCAGTCTGTAGCACTGTGTTTGACAGGCCAGAT), DAR40TT $\alpha$  (5'-CTGCAGGGTTTTTGTTCAGTCTGTAGCACTGTG-TGACAGGCCAGATC), DAR40TT $\beta$  (5'-CTGCAGGGTTTTTGTTCAGTCTGTAGCACTGTGT), DG62ABA-1 (5'-CTGCAGGGTTTTTGTACAGCCAGACAGTGGAGTACTACCACTGTG[aba]AAGACAGGCCAGATC), and DG62ABA+2 (5'-CTGCAGGGTTTTTGTACAGCCA GACAGTGGAGTACTACCACTGTG[aba]GTAAGACAGGCCAGATC). The standard uncleaved 12 and 23 RSS substrates were made by annealing DAR39/40 and DG61/62, respectively. In the prenicked versions, DAR39 was replaced by DAR42 and DG10, while DG61 was replaced by DAR42 and DG4. For 3'-end labeling with Klenow enzyme, the standard oligonucleotides were truncated by one residue to provide a fill-in site. To change the standard coding flank from 5'-TAA to 5'-ATA, DAR40TTT was annealed to DAR39AAA or DAR42AAA.

The 23 RSSs with abasic residues (shown by "aba" in brackets in the above sequences) were made by annealing DG62ABA-1 or DG62ABA+2 with DG61, DAR42, and DG4 as required. Thymine dimer-modified substrates were made by annealing DAR40TT $\alpha$  or DAR40TT $\beta$  with DAR39AAA for the uncleaved RSS or with DAR42AAA and DG10 for the prenicked RSS.

All oligonucleotides were polyacrylamide gel electrophoresis (PAGE) purified as single-stranded DNA before use. If required, they were 5'-end labeled with polynucleotide kinase. They were annealed by heating to 96°C for 1 min in 20 mM HEPES (pH 7.0)–100 mM potassium glutamate followed by slowly cooling to room temperature. Radiolabeled oligonucleotides (25 pmol in 50  $\mu$ l) were annealed with a 50% excess of cold partners. Double-stranded oligonucleotides were PAGE purified before use.

Substrates for cross-linking were internally labeled with a single  $^{32}$ P atom immediately 5' or 3' of the IdU substitution (given in brackets in the sequences above) (25). QME27 and QME28 were 5'-end labeled using polynucleotide kinase and individually ligated to QME29 using OLIGO-RAG as a complementary template scaffold. The full-length ligated products, QME27-29 and QME28-29, were purified by denaturing PAGE and individually annealed to WTOP to provide uncleaved IdU-substituted 12 RSS substrates. For the prenicked substrates, WTOP was replaced with the TOPARM and TOPFLANK oligonucleotides.

**Proteins.** Murine RAG1 and -2 were expressed as fusions with glutathione S-transferase (GST) and purified from CHO cell culture as described previously (45). RAG1 core (positions 384 to 1008) was either wild type or had point mutations as specified in the text. RAG2 was the wild-type core sequence (positions 1 to 387). HMGB1 was purified as described previously (44).

**Recombination reactions, EMSA, and KMnO<sub>4</sub> sensitivity.** Each 40- $\mu$ l reaction mixture contained 25 mM HEPES (pH 7.0), 50 mM K glutamate, 0.1 mg/ml bovine serum albumin, 8% glycerol, and 0.4 pmol radiolabeled DNA, plus 1  $\mu$ l of RAG1 or -2 (~75 ng each, at 0.75 pmol and 1.1 pmol, respectively). Coupled 12/23 reaction mixtures contained 0.08 pmol of radiolabeled partner plus 3.2 pmol of the unlabeled partner. Coupled reaction mixtures in addition contained 100 ng (4 pmol) of HMGB1 protein. The following additions were made as indicated: 2 mM CaCl<sub>2</sub>, 4 mM MgCl<sub>2</sub>, and 4 mM EDTA. After 5 min at room temperature (~25°C), 2.5  $\mu$ g of sheared salmon sperm DNA was added and the mixture was incubated for a further 10 min. To check the extent of complex formation, 10  $\mu$ l was supplemented with 2  $\mu$ l 50% glycerol and analyzed by electrophoretic mobility shift assay (EMSA) using a 4% acrylamide-bis (29:0.8) gel in 0.5 $\times$  Tris-borate-EDTA buffer. Electrophoresis was for 90 min at 25 mA at 4°C. The remaining 30  $\mu$ l of the reaction mixture was treated with KMnO<sub>4</sub> for 30 s and analyzed on a DNA sequencing gel as described previously (7).

**Protein-DNA cross-linking.** The cross-linking strategy was similar to that described by Eastman and colleagues (25). Complexes were assembled as described above. Reactions were placed as a drop on a sheet of Parafilm and were exposed to a Stratagene transilluminator (312-nm tubes) from above for 13 min at a distance of 4 cm. Reaction mixtures were recovered into tubes, heated to 68°C for 10 min to denature the RAG proteins, and adjusted to 10 mM MgCl<sub>2</sub> and 1 mM CaCl<sub>2</sub>. DNase I (2  $\mu$ g) and micrococcal nuclease (0.2  $\mu$ g) were added, and the mixture was incubated for 1 h at 37°C. Reaction mixtures were mixed with Laemmli loading buffer supplemented with 1% sodium dodecyl sulfate (SDS), heated for 10 min at 95°C, and analyzed by SDS-PAGE (7.5% acrylamide) with a stacking gel containing 1% SDS. The gel was dried and recorded by autoradiography. Prestained markers (NEB) were used to locate cross-linked products.

## RESULTS

**Experimental system.** V(D)J recombination *in vivo* is governed by the 12/23 rule, as originally suggested by Early and colleagues (23). However, several exceptions to this rule have been documented: for example, 23/23 recombination detected in a mouse strain with a noncanonical immunoglobulin H locus (38) and the apparent cleavage of isolated 12 RSS sequences (49, 63). The 12/23 rule also operates *in vitro* in reactions using purified components (24, 36, 60, 72, 73). However, under these artificial conditions it is probably less well enforced, and cleavage is readily detected in reactions containing an RSS of one type. One investigation into the requirements for cleavage using maltose-binding protein-RAG fusion proteins concluded that there is a >15-fold preference for the 12/23 reac-

tion over a 12/12 reaction and that this preference is enforced about equally by preferential synapsis and by control of the chemical steps required for hairpin formation (66, 67). The particular reaction conditions also affect the sway of the 12/23 rule. Cleavage of the 12 RSS alone can be increased significantly by the method of RAG protein purification, by extended incubation times, and by using Mn<sup>2+</sup> in place of Mg<sup>2+</sup>, which is the physiological divalent metal ion.

In this work, we use the GST-RAG fusion proteins which are probably less tightly governed by the 12/23 rule than the corresponding maltose-binding protein-RAG fusion proteins. Cleavage of the 12 RSS in the presence of Mg<sup>2+</sup> is quite efficient, as previously noted (58). Nevertheless, our reactions conform to the 12/23 rule when both RSSs are present in the reaction mixture prior to addition of the RAG proteins. When coupled 12/23 reactions were required, we added the 23 RSS in threefold molar excess to reinforce the natural preference for the 12/23 synapsis (Materials and Methods).

**Permanganate sensitivity of the 12 RSS.** In the Tn5 transpososome, the T+2 base on the second strand is rotated almost 180° out of the helix during the cleavage step of the reaction (Fig. 1). However, distortion of the helix can be detected even before the start of the reaction by a mild permanganate sensitivity of thymidine residues in the uncleaved complex. Following the first nick, permanganate sensitivity of these residues increases markedly when base flipping opens the DNA helix in preparation for the hairpin step (7).

To investigate whether RAG employs base flipping, we began by probing the permanganate sensitivity of the second strand (Fig. 2a). Complexes were formed in the presence or absence of divalent metal ion, as indicated, and treated briefly with permanganate. There was a slight, but reproducible, RAG-dependent increase in the sensitivity of the thymidine at position +10 of the second strand (henceforth "T+10"). This indicates that the RAG proteins bind to the 12 RSS in the absence of metal ions, even though such complexes are barely detectable in an EMSA. With the uncleaved substrate (left panel), Ca<sup>2+</sup> caused a very slight, but reproducible, increase in the sensitivity at T+2 and T-1.

When the experiment was performed with a prenicked substrate, the mild sensitivity of residues T+10 and T+2 to permanganate was similar to the uncleaved substrate (Fig. 2a, right panel). However, the sensitivity of T-1 was greatly increased in the presence of Ca<sup>2+</sup> which supports stable complex formation, but not catalysis. The sensitivity of T-1 indicates a significant distortion of the DNA following the first nick, as has been suggested by others on the basis of similar evidence (2, 12, 50, 59, 68). If the distortion at T-1 equates to V(D)J base flipping, then its location in the coding flank mirrors the reversed hairpin polarity of the hairpin compared to Tn5.

If the 12 RSS is incubated with Mg<sup>2+</sup> instead of Ca<sup>2+</sup>, cleavage takes place as the prenicked substrate is converted to the hairpin product (Fig. 2a, lane 9). Permanganate sensitivity at T-1 is lost, showing that this base has regained the protective stacking interactions with adjacent residues in the hairpin product. This may seem counterintuitive. However, within the loop region of a hairpin, which is generally between two and four residues long, depending on the sequence, the bases can be stacked and may also engage in various noncanonical base-pairing interactions (e.g., see references 11 and 70 and the





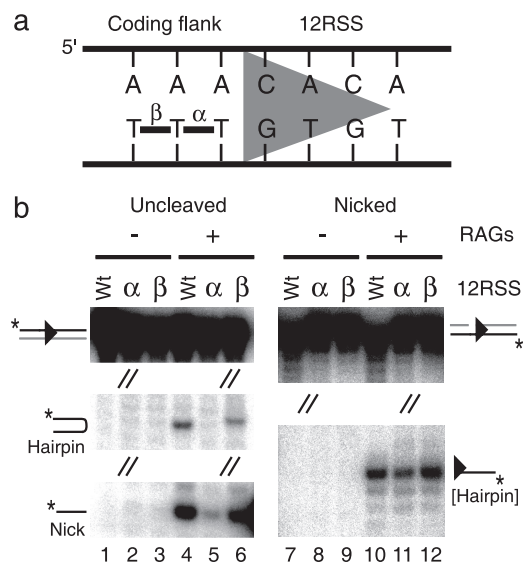


FIG. 3. Thymine dimer tethering. Thymine dimer intrastrand cross-links were incorporated into the bottom strand of the 12 RSS during oligonucleotide synthesis. (a) The locations of the thymine dimer tethers are illustrated. The  $\alpha$ -tether is between T-1 and T-2, while the  $\beta$ -tether is between T-2 and T-3. (b) RAG cleavage reactions were performed with uncleaved and prenicked substrates. An autoradiogram of a denaturing DNA sequencing gel is shown. The diagonal slashes indicate regions of the gels that have been removed because they contain no relevant information. Uncropped images of the gels are provided in Fig. S2 in the supplemental material. The location of the radioactive label is indicated by an asterisk. The strand illustrated in gray is unlabeled and therefore invisible. Wt, wild type.

RAG complex that moves the IdU residue within range of a cross-linkable group such as an aromatic amino acid side chain. In the case of Tn5, the site of T+2 cross-linking is the stacking tryptophan residue (7).

**Covalent tethering of T-1 and T-2.** The permanganate sensitivity and cross-linking assays suggested that the T-1 residue in the 12 RSS might be the equivalent of the flipped base in Tn5. To further test this hypothesis, we designed a modified 12 RSS substrates in which T-1 and T-2 residues were covalently linked by a *cis-syn* thymine dimer incorporated during oligonucleotide synthesis. We refer to this as the  $\alpha$ -modification (Fig. 3a). We reasoned that the rigidity of this molecular tether would prevent the rotation of one residue relative to the other. Furthermore, the bulk of the thymine dimer would prevent the insertion of T-1 into a close-fitting binding pocket in RAG1, analogous to that which accommodates the flipped base in Tn5. We also prepared a second substrate, the  $\beta$ -modification, in which residues T-2 and T-3 were linked.

The thymine dimer-modified 12 RSS substrates were tested in 12/23 cleavage reactions (Fig. 3b). In this experiment, the modified 12 RSS is radioactively labeled and the reaction is supplemented with a threefold molar excess of unlabeled 23 RSS. The  $\alpha$ -tether between T-1 and T-2 inhibited the nicking step severely (lane 5) but was much less detrimental to hairpin formation with a prenicked substrate (lane 11). The  $\beta$ -tether between T-2 and T-3 did not affect either step of the reaction significantly (lanes 6 and 12). One potential caveat to any interpretation of this experiment is that the thymine dimer

modification causes a kink in the DNA (52). The  $\alpha$ -tether between T-1 and T-2 may therefore be unfavorable to the nicking step for reasons unconnected to base flipping. Nevertheless, the insensitivity of the reaction to the  $\beta$ -tether supports the notion that steric freedom at T-1 is particularly important. Finally, the significant hairpinning activity with the prenicked  $\alpha$ -tether suggests that the T-1 base need not enter a highly specific binding pocket in the protein during this step. It appears that RAG and Tn5 transposase are significantly different in this respect. However, this is not completely unexpected. The -1 position of the V(D)J coding flank can be occupied by any one of four bases, whereas the flipped base in Tn5 and Tn10 is always a thymidine and can therefore be accommodated in a specific binding pocket. If the coding flank T-1 is indeed flipped, as it appears to be from the evidence already presented, perhaps it is accommodated in a nonspecific cleft, analogous to that seen in the structures of other enzymes that are able to flip different types of bases (Discussion).

**Conformational coupling in the 12/23 RSS complex.** In the presence of the HMG cofactor, assembly of the mixed 12/23 RSS complex is strongly favored over either of the two unmixed complexes, particularly the 23/23 pair (described above and see references 24, 29, 36, 60, and 73). This bias allows opposite sides of the paired complex to be experimentally distinguished simply by labeling only one of the two partners.

Very little is known about the structure of the synapsis or the identity of the residues in the DNA or protein that mediate the 12/23 rule. We therefore wished to ask whether the conformational change at T-1 is coupled on either side of the RAG complex. RAG complexes were assembled in the presence of  $\text{Ca}^{2+}$  ions using a radioactively labeled prenicked 12 RSS (Fig. 4a). As demonstrated above, residue T-1 was sensitive to permanganate oxidation (lane 4). The sensitivity was unaffected by the addition of HMG protein (lane 5). However, when an uncleaved 23 RSS partner was provided in the reaction mixture during assembly of the complexes, permanganate sensitivity at T-1 was lost (lane 6). In contrast, when a prenicked 23 RSS was provided, the permanganate sensitivity was unaltered (lane 7). The permanganate sensitivity of the nicked 12 RSS is thus dependent on the state of the partner in the 12/23 synapsis.

The reciprocal experiment was also performed by assembling RAG complexes with a radioactively labeled prenicked 23 RSS (Fig. 4b). The permanganate sensitivity of T-1 was similar to that in the 12 RSS, except that it was dependent on the HMGB1 cofactor required for RAG binding to the 23 RSS (compare lanes 4, 11, and 12). The permanganate sensitivity of the 23 RSS was abolished when an uncleaved 12 RSS partner was provided (lane 13), but not when the partner was prenicked (lane 14). It therefore appears that the distortion at T-1 is coupled on either side of the complex and can take place only after both sides of the complex have been nicked. Since it seems likely that this conformational change is a prerequisite for the hairpin step, we will, for convenience, sometimes refer to it as an "activation" step.

**Transduction of the activation signal requires T+2 but not T-1.** We wondered how the RAG proteins sensed nicking of the RSSs and how the signal was transmitted to the other side of the complex. For example, we wanted to determine whether perhaps the movement of T-1 provided a direct cue for the

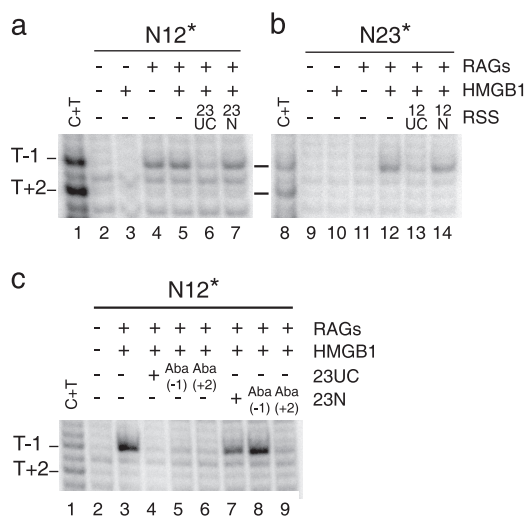


FIG. 4. Conformational coupling in the RAG complex. RAG complexes were assembled in the presence of  $\text{Ca}^{2+}$  and treated with permanganate as described in the legend to Fig. 2. The substrates were likewise labeled on the 5' end of the second strand. Autoradiograms of denaturing DNA sequencing gels are shown. Two aspects of the in vitro reaction conditions used here could potentially bias the results. The RSS were assembled using synthetic oligonucleotides, and there was no 5' phosphate group adjacent to the nick in the prenicked substrates. Furthermore, although  $\text{Ca}^{2+}$  is widely used as a noncatalytic analog of the catalytic  $\text{Mg}^{2+}$ , it is probably coordinated atypically in the active site. We have therefore performed control experiments to confirm that neither of these conditions is responsible for the permanganate sensitivity at T-1 (see Fig. S4 in the supplemental material). N, prenicked; UC, uncleaved; Aba, abasic residue at specified position. (a) Reaction mixtures contained a labeled prenicked 12 RSS (N12\*). Mixtures were supplemented with unlabeled 23 RSS partners, which were either uncleaved or prenicked, as indicated. (b) The reciprocal experiment to that shown in panel A. A labeled prenicked 23 RSS (N23\*) was supplemented with unlabeled 12 RSS partners. (c) A similar experiment to that shown in panel a. A labeled prenicked 12 RSS was combined with 23 RSS partners that contained abasic residues at positions -1 and +2 of the second strand.

activation step. Alternatively, the nick may be sensed at another site by a change in the curvature, stiffness, or tension in the DNA, for example.

We assembled RAG complexes in which a nicked 12 RSS was paired with various 23 RSS partners that were either normal or incorporated abasic residues at the T-1 and T+2 positions (Fig. 4c). The abasic sites were incorporated during oligonucleotide synthesis using the dSpacer modification that provides a normal sugar-phosphate backbone with a missing base. As before, the permanganate sensitivity of the nicked 12 RSS was lost when an uncleaved 23 RSS partner was provided (compare lanes 3 and 4). This behavior was unchanged by the presence of abasic sites at positions -1 and +2 (lanes 5 and 6). This shows that the uncleaved status of the partner RSS is not sensed directly by the presence of either of these residues in the helix.

We next wished to ask whether either of these residues signaled the nicked status of a nicked partner RSS. When a prenicked 23 RSS partner was provided, the permanganate sensitivity of the 12 RSS was maintained, as expected (lane 7). This behavior was unchanged by the abasic modification at position -1 (lane 8). The T-1 base is therefore not directly

involved in signaling the nicked status of the RSS. However, an abasic residue at position +2 abolished permanganate sensitivity and the prenicked 23 RSS behaved as if it was uncleaved (lane 9). This result shows that T+2 is critical for transduction of the nicking signal, but it is of course unclear whether this is a direct or indirect effect.

#### Hairpin-defective RAG1 mutants block the activation step.

Two groups have searched for aromatic amino acid residues in RAG1 that might perform the same function as the tryptophan residue in Tn5 transposase that stacks the flipped T+2 base (27, 45). Both groups searched for mutants that were nicking proficient but hairpin defective. Grundy and colleagues proposed W893 and W965 as likely candidates (27). In their conclusion, however, they slightly favored W956 because it was rescued by an abasic substitution at T-1 and because the hairpin step with the W893 mutation was rescued by  $\text{Mn}^{2+}$  and certain coding flank sequences. Lu and colleagues tested every aromatic residue in RAG1 and identified two additional residues, Y935 and F971, as candidates (45). However, they assessed the mutations using slightly different criteria and concluded that W893 was the best candidate stacking residue. Lu et al. eliminated Y935 as a candidate by showing that it was the hydroxyl group rather than the aromatic ring that was important for hairpinning. They also eliminated the W956 residue on the grounds that the W956A mutation was primarily a nicking mutant. We investigated this point further and found that the behavior of the W956A RAG1 mutant is complex (see Fig. S3 in the supplemental material). The W956A mutation is almost completely defective for nicking and hairpin formation with the 12 RSS alone, irrespective of whether it is uncleaved or prenicked. However, it provides about 30% hairpin formation, compared to the wild type, in 12/23 reactions provided that both RSS are prenicked (see Fig. S3, lane 8, in the supplemental material). A similar effect was noted by Grundy et al. (27) using  $\text{Mn}^{2+}$  in the reaction.

The four RAG1 mutations discussed above were next tested for their effect on the permanganate sensitivity of the coding flank T-1 residue (Fig. 5a). In our standard reaction with a prenicked 12 RSS, all four mutants, including W956A, were defective for the hairpin step, as indicated below the gel. There was a perfect correlation between the hairpin defect and the conformational change required for permanganate sensitivity at T-1. The alanine substitutions of all four aromatic residues were permanganate resistant (lanes 3, 4, 5, and 8). However, the serine and threonine substitutions of Y935, which provide a replacement hydroxyl group and remain hairpin proficient, were permanganate sensitive (lanes 11 and 12). In contrast, the hairpin-defective leucine and phenylalanine substitutions of Y935, which preserve the original residue's chain length and aromatic characters, respectively, were permanganate resistant (lanes 13 and 14). It seems that all of the hairpin-defective mutants have lost the ability to distort the DNA at the T-1 position. T-1 distortion is therefore probably a prerequisite for the hairpin step and may correspond to base flipping in Tn5 transposition.

We also tested two DDE mutants. Permanganate sensitivity was lost in the triple mutant, where the DDE residues had been changed to NNQ (Fig. 5A, lane 18). In contrast, the single E962A mutant remained permanganate sensitive (lane 16). Crystal structures for DDE superfamily members suggest



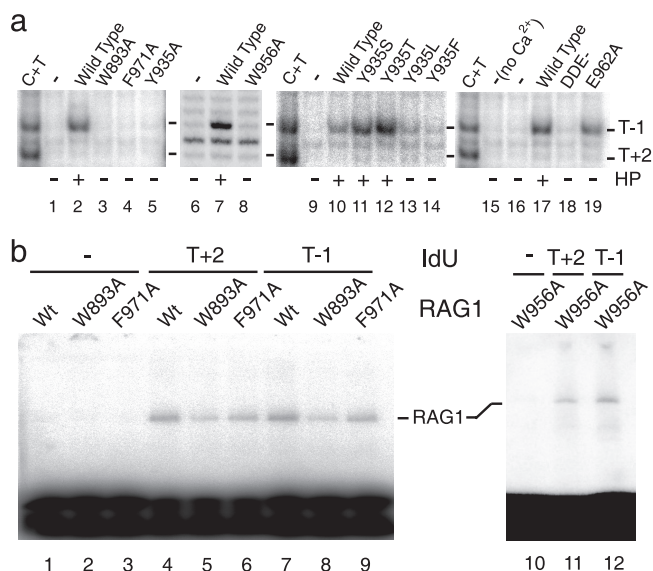


FIG. 5. Analysis of hairpin-defective mutations in RAG1. (a) RAG complexes were assembled on a prenicked 12 RSS in the presence of  $\text{Ca}^{2+}$  and treated with permanganate as described in the legend to Fig. 2. The wild-type (Wt) and mutant proteins formed approximately equal amounts of complex as determined using an EMSA (not shown). DDE<sup>-</sup> indicates a protein with an NNQ triple substitution. The hairpin proficiency of the respective proteins after the addition of  $\text{Mg}^{2+}$  is indicated by a + or - sign below each lane of the gel. (b) RAG complexes were assembled on a 12 RSS substituted with the IdU photo-cross-linking reagent at positions T+2 and T-1. Cross-linking was performed as described in Fig. 2c and Materials and Methods. An autoradiogram of a Laemmli SDS-PAGE is shown.

that the catalytic metal ions contribute significantly to positioning the substrate in the active site (51, 75). The loss of permanganate sensitivity in the triple mutant probably reflects the loss of both divalent metal ions from the active site, in contrast to the probable loss of just one metal ion in the case of the single mutant.

**T-1 and T+2 cross-linking of RAG mutants.** In Tn5 transposase, mutation of the stacking tryptophan residue abolishes photo-cross-linking of the transposase to the flipped base (7). We therefore tested whether the remaining candidate stacking-residues, W893, W956, and F971, were required for RAG1 cross-linking by IdU substitution at positions -1 and +2 in a prenicked 12 RSS (Fig. 5b). Cross-linking was not abolished by the alanine substitutions at any of these positions. Instead, each mutation reduced cross-linking at each position by a similar amount. This is evidence against a direct stacking interaction between either of these bases and any of the amino acid residues. However, the residual cross-linking of the mutant RAG proteins on the prenicked 12 RSS suggests that they have achieved a conformational change, compared to the uncleaved situation (Fig. 2c, lanes 5 and 9), but that it falls short of that required for the distortion of the T-1 residue detected by permanganate sensitivity.

## DISCUSSION

**The relationship between the proximal and distal hairpin mechanisms.** The polarity of the phosphoryl transfer reactions

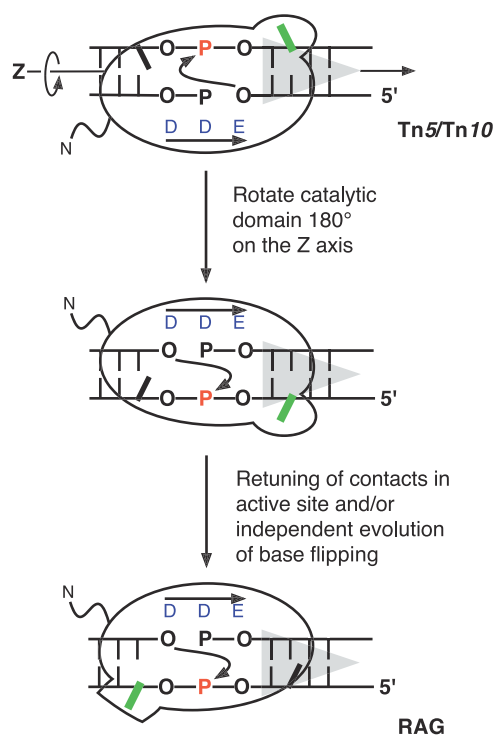


FIG. 6. A model for reversal of the hairpin polarity. The hairpin step of the reaction is illustrated with the nucleophilic 3'-OH group attacking the scissile phosphate on the second strand. For simplicity, the respective recombinases are arbitrarily illustrated acting in *cis*. In Tn5 (top), the flipped thymine base at position +2 is accommodated in a specific binding pocket in the transposase, where it is stacked on the W298 residue (see text for details). In Tn5 (top), the base at position -1 of the second strand is displaced to indicate its mild permanganate sensitivity. In two dimensions, the translocation of the active site from one strand to the other is most easily illustrated by a 180° rotation about the z axis of the DNA. If the position of the flipped base was maintained relative to other structural features of the protein, it would now be located on the transferred strand, 5' to the scissile phosphate. However, we propose a different model in which the flipped base in RAG recombination is at position T-1, where it maintains its position 3' to the scissile phosphate (bottom). Furthermore, the flipped base, which can be purine or pyrimidine owing to its location in the coding flank, is accommodated in a nonspecific binding pocket within the recombinase. Further details of the model are given in the text. Flipped bases are indicated above and below the helix. In the RSS (bottom), the base at position +2 of the second strand is displaced to indicate its role in signaling the nicked status of the RSS (Fig. 4c and see the text). The model implies that the putatively flipped base at position -1 of the RSS (bottom) is equivalent to the distorted base at position -1 of Tn5 (top).

in transposition and V(D)J recombination are defined by the asymmetric location of the recombinase recognition sequences on one side of the recombination site (Fig. 6). We will refer to the different hairpin reactions as proximal (Tn5) or distal (RAG and hAT) to the respective binding sites. The most parsimonious way to switch between the proximal and distal hairpin mechanisms is for the recombinase active site to translocate between opposite strands across the major or the minor groove (51, 75). In two dimensions, this movement round the helix can be illustrated as a 180° rotation of the catalytic domain about the z axis of the DNA (Fig. 6, top and middle elements). This reorientation would have to be accompanied

by retuning the protein-DNA contacts to accommodate the reversed polarity of the DNA strand. The relative positions of the DDE residues would also change with respect to the various chemical groups on the DNA. This is most easily appreciated by noticing how the first D residue exchanges position with the E residue above the nucleophilic oxygen atom. However, changing the relative positions of catalytic residues may not be as problematic as it might seem at first sight (51, 75). In the Tn5 transpososome, at least, the metal ions appear to be coordinated symmetrically, suggesting that the active site should accommodate strands of either polarity with minimal tuning required. More generally, the two-metal-ion mechanism appears to be versatile, and distantly related members of the family appear to accommodate DNA strands of opposite polarity (51, 75).

The flipped base, which lacks the structural symmetry of the active-site metal ions, would place a more severe constraint on the reversal of the hairpin polarity. Rotation of the active site, as it moved from one strand to the other, would place the hairpin binding pocket on the 5' side of the scissile phosphate, instead of the 3' side (Fig. 6, middle element). It seems likely that this location would affect its functions. In the Tn5 transpososome, the flipped base is important for all four chemical steps of the reaction. It is particularly important for hairpin resolution and the integration steps, when it appears to hold the second strand in position (3, 7, 37). Of course, neither of these steps is performed by the RAG proteins, which are only concerned with the nick and hairpin steps of the reactions. Even so, the new position of the flipped base on the transferred strand would represent a significant mechanistic innovation for *Transib* and the hAT transposons, which must still perform integration.

**Assessment of RAG1 candidate stacking residues.** The most significant similarity between the Tn5 and the RAG recombination reactions is a conformational change following the first nick. In Tn5, it is detected as the permanganate sensitivity of the second base 3' to the scissile phosphate (T+2). In RAG recombination, the first base 3' to the scissile phosphate is permanganate sensitive (T-1). In the Tn5 transpososome, the conformational change also promotes photo-cross-linking between T+2 and the transposase. Furthermore, the cross-linking is abolished by mutating the stacking tryptophan residue to alanine. However, although alanine substitution of the candidate aromatic residues in RAG1 abolished permanganate sensitivity, cross-linking at positions T+2 and T-1 of the RSS was reduced but not eliminated (Fig. 5). This suggests that although the aromatic residues are required for the conformational change, they do not directly contact either of the bases.

The W956 candidate stacking residue is located just before the residue of the RAG1 DDE catalytic triad. This suggests that it is more likely to be equivalent to the Tn5 transposase probe residue (W323) that occupies a similar position and helps force the flipped base from the helix. Indeed, this possibility was mentioned by Grundy et al. (6, 7, 27). None of the present results are inconsistent with this proposition.

In the Tn10 transpososome, the sensitivity of the flipped base at T+2 is comparable to the sensitivity of T-1 in the 12 RSS (6). However, in Tn5 the flipped base is protected by the stacking tryptophan and/or its location within a hydrophobic binding pocket (7). Its extrahelical conformation is revealed by

mutating the tryptophan residue to alanine. It is therefore possible that the flipped base in V(D)J recombination is really at position T+2 of the RSS, but that its extrahelical conformation is masked by stacking on an aromatic amino acid side chain or sequestration within a hydrophobic pocket in the protein. However, none of the mutations in the candidate residues revealed permanganate sensitivity of this residue, as might be expected if this was the case (Fig. 5a). We therefore conclude that none of the RAG1 candidate residues are likely to perform a function equivalent to the stacking tryptophan residue in Tn5 transposase.

**A new but familiar position for the flipped base, 3' to the scissile phosphate.** We would like to suggest that it is the base at position -1 on the second strand of the RSS that is flipped from the helix (Fig. 6, bottom element). The positions of the flipped bases in Tn5 and the RSS, just 3' to the respective scissile phosphates, are consistent with their functional equivalence. However, in this model the flipped bases do not interact with homologous regions of the respective recombinases (Fig. 6, compare top and bottom elements). The respective recombinases both make important interactions with T+2 on the respective second strands. However, the difference is the position of the flipped base, which, unlike the catalytic metal ions, is intrinsically asymmetric and cannot therefore interact with the same region of the protein after the polarity of the reaction chemistry has been reversed.

Base flipping is a common theme in nucleic acid metabolism. It has evolved several times, and several distinct mechanisms have been documented (56, 62, 64, 65). Even if the flipped base in Tn5 represents an ancestral condition for all cut-and-paste transposons, there is no reason to suppose that the stacking interactions would have been conserved when the polarity of the reaction chemistry was reversed. RAG may instead resemble the T4 pyrimidine dimer DNA glycosylase, which Stivers refers to as an "atypical" enzyme in which the flipped base is held in a featureless protein crevice (65). This would be consistent with the cross-linking data and the assessment of the candidate stacking residues. Indeed, the -1 position in the coding flank can be occupied by purine or pyrimidine bases that would be difficult to accommodate in a highly specific binding pocket. These ideas are also consistent with numerous examples of "good" and "bad" coding flank sequences and their rescue by mispaired bases (e.g., see references 17, 54, and 74).

Finally, in our view it seems plausible that the distal hairpin mechanism, exemplified by RAG and the hAT family of transposons, evolved independently from an RNase H-like ancestor with a single-strand nicking activity. Some support for this idea is provided by a comparison of the Tn5 and *hermes* (hAT) transposase crystal structures (19, 28). In Tn5 transposase, the RNase H structural fold is interrupted by the insertion of an extra subdomain that provides the stacking tryptophan residue and the hydrophobic pocket that stabilizes the flipped base. The RNase H fold of *hermes* transposase is also disrupted by the insertion of an additional subdomain. The interruption is at the same point as in Tn5 transposase, but the inserted domain is structurally unrelated. The common point of insertion could be a coincidence. However, it seems likely that the inserted subdomain in *hermes* transposase performs a function related to the hairpin reaction. The absence of homology between the



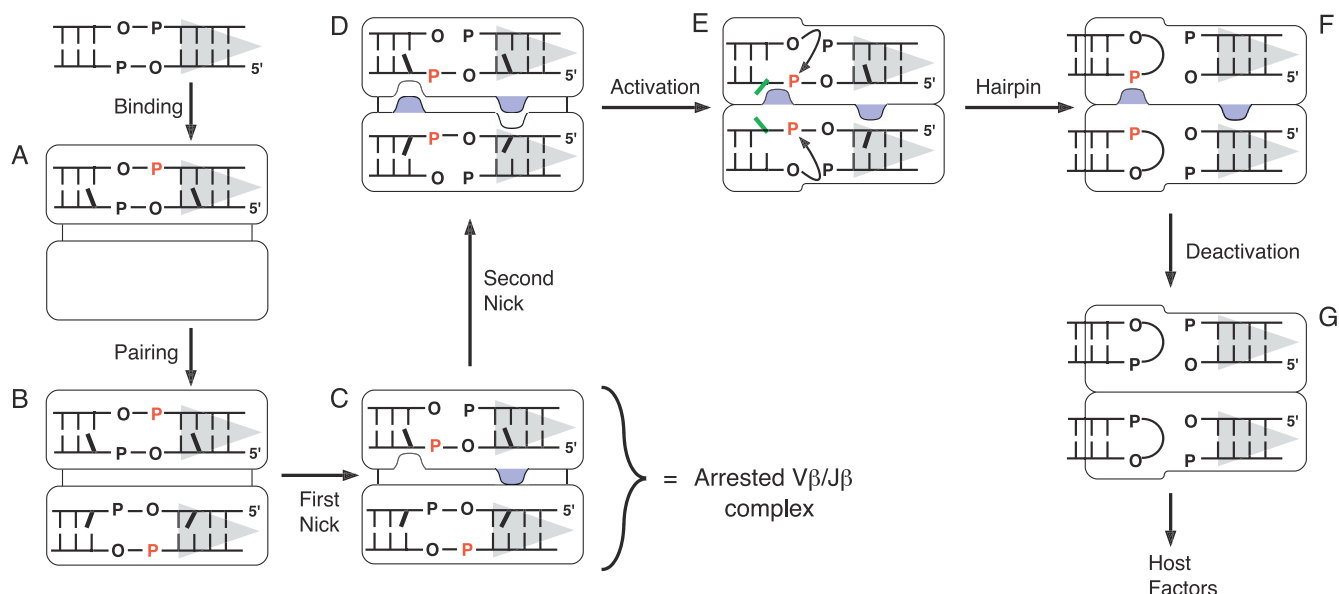


FIG. 7. Conformational coupling in the RAG synaptic complex. A dimer of RAG1/2 is represented by a pair of linked rectangles. The scissile phosphates are shown in red. RAG binds an RSS and recruits a second RSS (complexes A and B). The slight permanganate sensitivity at T+2 and T-1 is indicated by the displacement of the bases. Pairing is followed by the first nick, which causes a conformational change on one side of the complex (complex C). When the second RSS achieves the first nick (complex D), there is a second conformational change, or "activation step," detected as permanganate sensitivity at position T-1 (complex E). This conformational change allows the hairpin step to take place and completes cleavage of the RSSs (complex F). Since the hairpin ends are retained in the complex for a time, a final "deactivation" step is probably required to prevent reversal of the hairpin step, which is in principle reversible (complex G). However, none of the experiments presented here address this final point, which is included simply to complete the reaction scheme. Complex C represents the point at which the V $\beta$ /J $\beta$  complexes are proposed to be arrested by the B12/23 restriction (see text for details). The unnicked partner in this complex represents the J $\beta$  RSS. It is of interest to note that cleavage in the Tn5 transpososome is coordinated by the *trans* architecture of the symmetrical complex. However, cleavage in the related Tn10/IS10 transpososome is modulated by additional factors. Like RAG, IS10 has an asymmetric synaptic complex with a host factor bound to one of the transposon ends. In IS10, these components mediate conformational coupling between opposite sides of the synaptic complex, affecting hairpin resolution and dictating the choice of recombining partners and target sites (6, 9, 15, 16, 42, 43, 61).

inserted subdomains in the *hermes* and Tn5 transposases, in our view, reflects the independent origins of the proximal and distal hairpin reactions.

**Coupled cleavage and control of the hairpin step.** Cleavage at an unsynapsed Tn5 end is mechanistically excluded by the monomeric state of the unbound transposase and the *trans* architecture of the synaptic complex. RAG1 also cleaves the RSS bound in *trans* using a single active site, but its initial binding as a dimer from solution allows the possibility of uncoordinated cleavage events (69). Indeed, RAG1 can probably nick an unsynapsed 12 RSS, but the hairpin step depends on synapsis and prior nicking of the partner end (21, 36, 73, 76). Very little is known of the mechanism that coordinates the hairpin step on either side of the complex. Our results suggest that the hairpin reactions are coordinated by conformational coupling of base flipping on either side of the complex. Together with a large body of preexisting knowledge, this is summarized by the model presented in Fig. 7.

The first step of the reaction is the interaction of an RSS with dimer of RAG1/2 (Fig. 7, complex A). A second unbound RSS is next recruited into a paired complex (complex B). This is followed by a nicking of one of the two partner RSSs (complex C). This causes a conformational change in one-half of the complex that signals the nicked status of the bound RSS. However, the uncleaved status of the partner RSS prevents progression to the base-flipping stage which is required for the

hairpin reaction. In our experiments, complex C is equivalent to those mixed complexes in which the T-1 base was resistant to permanganate treatment (e.g., the N12/UC23 complex in Fig. 4c, lane 4). Nicking at the second RSS promotes a matching conformational change in the other half of the complex (complex D). Both halves of the paired complex are now signaling the nicked status of the respective RSSs. There follows a second conformational change, indicated as the "activation" step, which is accompanied by base flipping at position T-1 (complex E). This conformational change affects both sides of the complex, and is detected in our experiments by the permanganate sensitivity at T-1. The complex is now proficient for the hairpin step (transesterification) which completes cleavage of the RSSs (complex F). Initially, the coding flanks are retained in the RAG complex. Since transesterification maintains the energy of the phosphodiester bond, a final conformational change is required to prevent the reversal of the reaction (complex G).

In Tn5, transposition-coupled cleavage is imposed by the monomeric state of the unbound transposase and the *trans* architecture of the synaptic complex. However, it seems inflexible compared to the conformational coupling and two-step activation of the RAG complexes in V(D)J recombination. Flexibility is important for the 12/23 rule and in the choice of recombining partners. For example, if RAG binding to a favorable RSS was stabilized by the one-sided conformational

change that accompanies the first nick, it would provide an extended window to scan for a second favorable RSS, minimizing the number of unproductive nicks created while searching for recombining partners.

The B12/23 rule describes another form of restriction that excludes direct V $\beta$ -to-J $\beta$  joining in the TCR  $\beta$  region (4, 21, 31, 32). Joining of these segments is inefficient despite the presence of appropriately oriented pairs of 12 and 23 RSS. It appears that this restriction is imposed equally by the spacer and the coding flank sequences, which together abolish nicking of some J $\beta$  segments almost entirely (21). Nevertheless, some of these slow nicking segments have been found to perform synapsis very well. Schatz and colleagues propose that dissociation of a V $\beta$ /J $\beta$  synaptic complex is faster than nicking of the J $\beta$  partner and that the nicked V $\beta$  substrate released from this complex can go on and recombine with a 5'D $\beta$  substrate (21). In our model, the arrested V $\beta$ /J $\beta$  synaptic complex would be represented by complex C. This illustrates the importance of conformational coupling in the RAG complex. Even though all of the protein and DNA components have been assembled together in a single complex, the first double-strand break at one RSS is suppressed by the unfavorable partner RSS. This may well provide for the integration of different signals and flexible responses in V(D)J recombination.

#### ACKNOWLEDGMENTS

This work was funded by grants from The Wellcome Trust to R.C. and from the NIH to D.B.R. (AI-36420).

We thank Y. Bigot and F. X. Barre for allowing J.B. to perform in their labs experiments requested in review. R.C. thanks Patrick Swanson for helpful discussions during the preparation of the manuscript.

#### REFERENCES

- Agrawal, A., Q. M. Eastman, and D. G. Schatz. 1998. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* **394**:744–751.
- Akamatsu, Y., and M. A. Oettinger. 1998. Distinct roles of RAG1 and RAG2 in binding the V(D)J recombination signal sequences. *Mol. Cell. Biol.* **18**:4670–4678.
- Ason, B., and W. S. Reznikoff. 2002. Mutational analysis of the base flipping event found in Tn5 transposition. *J. Biol. Chem.* **277**:11284–11291.
- Bassing, C. H., F. W. Alt, M. M. Hughes, M. D'Auteuil, T. D. Wehrly, B. B. Woodman, F. Gartner, J. M. White, L. Davidson, and B. P. Sleckman. 2000. Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. *Nature* **405**:583–586.
- Bhasin, A., I. Y. Goryshin, and W. S. Reznikoff. 1999. Hairpin formation in Tn5 transposition. *J. Biol. Chem.* **274**:37021–37029.
- Bischoeur, J., and R. Chalmers. 2009. Base flipping in Tn10 transposition: an active flip and capture mechanism. *PLoS One* **4**:e6201.
- Bischoeur, J., and R. Chalmers. 2007. Base-flipping dynamics in a DNA hairpin processing reaction. *Nucleic Acids Res.* **35**:2584–2595.
- Bolland, S., and N. Kleckner. 1995. The two single-strand cleavages at each end of Tn10 occur in a specific order during transposition. *Proc. Natl. Acad. Sci. USA* **92**:7814–7818.
- Chalmers, R., A. Guhathakurta, H. Benjamin, and N. Kleckner. 1998. IHF modulation of Tn10 transposition: sensory transduction of supercoiling status via a proposed protein/DNA molecular spring. *Cell* **93**:897–908.
- Chalmers, R. M., and N. Kleckner. 1994. Tn10/IS10 transposase purification, activation, and in vitro reaction. *J. Biol. Chem.* **269**:8029–8035.
- Chou, S. H., K. H. Chin, and A. H. Wang. 2003. Unusual DNA duplex and hairpin motifs. *Nucleic Acids Res.* **31**:2461–2474.
- Ciubotaru, M., and D. G. Schatz. 2004. Synapsis of recombination signal sequences located in *cis* and DNA underwinding in V(D)J recombination. *Mol. Cell. Biol.* **24**:8727–8744.
- Claeys Bouaert, C., and R. Chalmers. Gene therapy vectors: the prospects and potentials of the cut-and-paste transposons. *Genetica*, in press.
- Coen, E. S., R. Carpenter, and C. Martin. 1986. Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus*. *Cell* **47**:285–296.
- Crellin, P., and R. Chalmers. 2001. Protein-DNA contacts and conformational changes in the Tn10 transpososome during assembly and activation for cleavage. *EMBO J.* **20**:3882–3891.
- Crellin, P., S. Sewitz, and R. Chalmers. 2004. DNA looping and catalysis: the IHF-folded arm of Tn10 promotes conformational changes and hairpin resolution. *Mol. Cell* **13**:537–547.
- Cuomo, C. A., C. L. Mundy, and M. A. Oettinger. 1996. DNA sequence and structure requirements for cleavage of V(D)J recombination signal sequences. *Mol. Cell. Biol.* **16**:5683–5690.
- Curcio, M. J., and K. M. Derbyshire. 2003. The outs and ins of transposition: from mu to kangaroo. *Nat. Rev. Mol. Cell Biol.* **4**:865–877.
- Davies, D. R., I. Y. Goryshin, W. S. Reznikoff, and I. Rayment. 2000. Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science* **289**:77–85.
- Dawson, A., and D. J. Finnegan. 2003. Excision of the *Drosophila* mariner transposon mos1. Comparison with bacterial transposition and v(d)j recombination. *Mol. Cell* **11**:225–235.
- Drejer-Teel, A. H., S. D. Fugmann, and D. G. Schatz. 2007. The beyond 12/23 restriction is imposed at the nicking and pairing steps of DNA cleavage during V(D)J recombination. *Mol. Cell. Biol.* **27**:6288–6299.
- Dyda, F., A. B. Hickman, T. M. Jenkins, A. Engelman, R. Craigie, and D. R. Davies. 1994. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **266**:1981–1986.
- Early, P., H. Huang, M. Davis, K. Calame, and L. Hood. 1980. An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. *Cell* **19**:981–992.
- Eastman, Q. M., T. M. Leu, and D. G. Schatz. 1996. Initiation of V(D)J recombination in vitro obeying the 12/23 rule. *Nature* **380**:85–88.
- Eastman, Q. M., I. J. Villey, and D. G. Schatz. 1999. Detection of RAG protein-V(D)J recombination signal interactions near the site of DNA cleavage by UV cross-linking. *Mol. Cell. Biol.* **19**:3788–3797.
- Fugmann, S. D., I. J. Villey, L. M. Ptaszek, and D. G. Schatz. 2000. Identification of two catalytic residues in RAG1 that define a single active site within the RAG1/RAG2 protein complex. *Mol. Cell* **5**:97–107.
- Grundy, G. J., J. E. Hesse, and M. Gellert. 2007. Requirements for DNA hairpin formation by RAG1/2. *Proc. Natl. Acad. Sci. USA* **104**:3078–3083.
- Hickman, A. B., Z. N. Perez, L. Zhou, P. Musigari, R. Ghirlando, J. E. Hinshaw, N. L. Craig, and F. Dyda. 2005. Molecular architecture of a eukaryotic DNA transposase. *Nat. Struct. Mol. Biol.* **12**:715–721.
- Hiom, K., and M. Gellert. 1998. Assembly of a 12/23 paired signal complex: a critical control point in V(D)J recombination. *Mol. Cell* **1**:1011–1019.
- Hiom, K., M. Melek, and M. Gellert. 1998. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell* **94**:463–470.
- Hughes, M. M., R. E. Tillman, T. D. Wehrly, J. M. White, and B. P. Sleckman. 2003. The B12/23 restriction is critically dependent on recombination signal nonamer and spacer sequences. *J. Immunol.* **171**:6604–6610.
- Jung, D., C. H. Bassing, S. D. Fugmann, H. L. Cheng, D. G. Schatz, and F. W. Alt. 2003. Extrachromosomal recombination substrates recapitulate beyond 12/23 restricted VDJ recombination in nonlymphoid cells. *Immunity* **18**:65–74.
- Kapitonov, V. V., and J. Jurka. 2005. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biol.* **3**:e181.
- Kennedy, A. K., A. Guhathakurta, N. Kleckner, and D. B. Haniford. 1998. Tn10 transposition via a DNA hairpin intermediate. *Cell* **95**:125–134.
- Kim, D. R., Y. Dai, C. L. Mundy, W. Yang, and M. A. Oettinger. 1999. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. *Genes Dev.* **13**:3070–3080.
- Kim, D. R., and M. A. Oettinger. 1998. Functional analysis of coordinated cleavage in V(D)J recombination. *Mol. Cell. Biol.* **18**:4679–4688.
- Klenchin, V. A., A. Czyz, I. Y. Goryshin, R. Gradman, S. Lovell, I. Rayment, and W. S. Reznikoff. 2008. Phosphate coordination and movement of DNA in the Tn5 synaptic complex: role of the (R)YREK motif. *Nucleic Acids Res.* **36**:5855–5862.
- Koralov, S. B., T. I. Novobrantseva, K. Hochedlinger, R. Jaenisch, and K. Rajewsky. 2005. Direct in vivo VH to JH rearrangement violating the 12/23 rule. *J. Exp. Med.* **201**:341–348.
- Landree, M. A., J. A. Wibbenmeyer, and D. B. Roth. 1999. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. *Genes Dev.* **13**:3059–3069.
- Lee, G. S., M. B. Neiditch, S. S. Salus, and D. B. Roth. 2004. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* **117**:171–184.
- Liu, D., J. Bischoeur, A. Siddique, N. Buisine, Y. Bigot, and R. Chalmers. 2007. The human SETMAR protein preserves most of the activities of the ancestral *Hsmar1* transposase. *Mol. Cell. Biol.* **27**:1125–1132.
- Liu, D., P. Crellin, and R. Chalmers. 2005. Cyclic changes in the affinity of protein-DNA interactions drive the progression and regulate the outcome of the Tn10 transposition reaction. *Nucleic Acids Res.* **33**:1982–1992.
- Liu, D., S. Sewitz, P. Crellin, and R. Chalmers. 2006. Functional coupling between the two active sites during Tn 10 transposition buffers the mutation of sequences critical for DNA hairpin processing. *Mol. Microbiol.* **62**:1522–1533.

44. Lu, C. P., J. E. Posey, and D. B. Roth. 2008. Understanding how the V(D)J recombinase catalyzes transesterification: distinctions between DNA cleavage and transposition. *Nucleic Acids Res.* **36**:2864–2873.
45. Lu, C. P., H. Sandoval, V. L. Brandt, P. A. Rice, and D. B. Roth. 2006. Amino acid residues in Rag1 crucial for DNA hairpin formation. *Nat. Struct. Mol. Biol.* **13**:1010–1015.
46. McBlane, J. F., D. C. van Gent, D. A. Ramsden, C. Romeo, C. A. Cuomo, M. Gellert, and M. A. Oettinger. 1995. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell* **83**:387–395.
47. Mitra, R., J. Fain-Thornton, and N. L. Craig. 2008. piggyBac can bypass DNA synthesis during cut and paste transposition. *EMBO J.* **27**:1097–1109.
48. Munoz-Lopez, M., A. Siddique, J. Bischerour, P. Lorite, R. Chalmers, and T. Palomeque. 2008. Transposition of Mboumar-9: identification of a new naturally active mariner-family transposon. *J. Mol. Biol.* **382**:567–572.
49. Nakajima, P. B., and M. J. Bosma. 1997. Characterization of excised DNA intermediates associated with V(D)J recombination at the T-cell receptor  $\delta$  locus. *Mol. Cell. Biol.* **17**:2631–2641.
50. Nishihara, T., F. Nagawa, T. Imai, and H. Sakano. 2008. RAG-heptamer interaction in the synaptic complex is a crucial biochemical checkpoint for the 12/23 recombination rule. *J. Biol. Chem.* **283**:4877–4885.
51. Nowotny, M., S. A. Gaidamakov, R. J. Crouch, and W. Yang. 2005. Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* **121**:1005–1016.
52. Park, H., K. Zhang, Y. Ren, S. Nadji, N. Sinha, J. S. Taylor, and C. Kang. 2002. Crystal structure of a DNA decamer containing a *cis-syn* thymine dimer. *Proc. Natl. Acad. Sci. USA* **99**:15965–15970.
53. Ramsden, D. A., and M. Gellert. 1995. Formation and resolution of double-strand break intermediates in V(D)J rearrangement. *Genes Dev.* **9**:2409–2420.
54. Ramsden, D. A., J. F. McBlane, D. C. van Gent, and M. Gellert. 1996. Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. *EMBO J.* **15**:3197–3206.
55. Rice, P., R. Craigie, and D. R. Davies. 1996. Retroviral integrases and their cousins. *Curr. Opin. Struct. Biol.* **6**:76–83.
56. Roberts, R. J., and X. Cheng. 1998. Base flipping. *Annu. Rev. Biochem.* **67**:181–198.
57. Roth, D. B., J. P. Menetski, P. B. Nakajima, M. J. Bosma, and M. Gellert. 1992. V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. *Cell* **70**:983–991.
58. Santagata, S., V. Aidinis, and E. Spanopoulou. 1998. The effect of Me2+ cofactors at the initial stages of V(D)J recombination. *J. Biol. Chem.* **273**:16325–16331.
59. Santagata, S., E. Besmer, A. Villa, F. Bozzi, J. S. Allingham, C. Sobacchi, D. B. Haniford, P. Vezzoni, M. C. Nussenzweig, Z. Q. Pan, and P. Cortes. 1999. The RAG1/RAG2 complex constitutes a 3' flap endonuclease: implications for junctional diversity in V(D)J and transpositional recombination. *Mol. Cell* **4**:935–947.
60. Sawchuk, D. J., F. Weis-Garcia, S. Malik, E. Besmer, M. Bustin, M. C. Nussenzweig, and P. Cortes. 1997. V(D)J recombination: modulation of RAG1 and RAG2 cleavage activity on 12/23 substrates by whole cell extract and DNA-bending proteins. *J. Exp. Med.* **185**:2025–2032.
61. Sewitz, S., P. Crellin, and R. Chalmers. 2003. The positive and negative regulation of Tn10 transposition by IHF is mediated by structurally asymmetric transposon arms. *Nucleic Acids Res.* **31**:5868–5876.
62. Slupphaug, G., C. D. Mol, B. Kavli, A. S. Arvai, H. E. Krokan, and J. A. Tainer. 1996. A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature* **384**:87–92.
63. Steen, S. B., L. Gomelsky, S. L. Speidel, and D. B. Roth. 1997. Initiation of V(D)J recombination in vivo: role of recombination signal sequences in formation of single and paired double-strand breaks. *EMBO J.* **16**:2656–2664.
64. Stivers, J. T. 2008. Extrahelical damaged base recognition by DNA glycosylase enzymes. *Chemistry* **14**:786–793.
65. Stivers, J. T. 2004. Site-specific DNA damage recognition by enzyme-induced base flipping. *Prog. Nucleic Acid Res. Mol. Biol.* **77**:37–65.
66. Swanson, P. C. 2002. A RAG-1/RAG-2 tetramer supports 12/23-regulated synapsis, cleavage, and transposition of V(D)J recombination signals. *Mol. Cell. Biol.* **22**:7790–7801.
67. Swanson, P. C. 2002. Fine structure and activity of discrete RAG-HMG complexes on V(D)J recombination signals. *Mol. Cell. Biol.* **22**:1340–1351.
68. Swanson, P. C. 2004. The bounty of RAGs: recombination signal complexes and reaction outcomes. *Immunol. Rev.* **200**:90–114.
69. Swanson, P. C. 2001. The DDE motif in RAG-1 is contributed in *trans* to a single active site that catalyzes the nicking and transesterification steps of V(D)J recombination. *Mol. Cell. Biol.* **21**:449–458.
70. van Dongen, M. J., M. M. Mooren, E. F. Willems, G. A. van der Marel, J. H. van Boom, S. S. Wijmenga, and C. W. Hilbers. 1997. Structural features of the DNA hairpin d(ATCCTA-GTTA-TAGGAT): formation of a G-A base pair in the loop. *Nucleic Acids Res.* **25**:1537–1547.
71. van Gent, D. C., J. F. McBlane, D. A. Ramsden, M. J. Sadofsky, J. E. Hesse, and M. Gellert. 1995. Initiation of V(D)J recombination in a cell-free system. *Cell* **81**:925–934.
72. van Gent, D. C., D. A. Ramsden, and M. Gellert. 1996. The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. *Cell* **85**:107–113.
73. West, R. B., and M. R. Lieber. 1998. The RAG-HMG1 complex enforces the 12/23 rule of V(D)J recombination specifically at the double-hairpin formation step. *Mol. Cell. Biol.* **18**:6408–6415.
74. Wong, S. Y., C. P. Lu, and D. B. Roth. 2008. A RAG1 mutation found in Omenn syndrome causes coding flank hypersensitivity: a novel mechanism for antigen receptor repertoire restriction. *J. Immunol.* **181**:4124–4130.
75. Yang, W., J. Y. Lee, and M. Nowotny. 2006. Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity. *Mol. Cell* **22**:5–13.
76. Yu, K., and M. R. Lieber. 2000. The nicking step in V(D)J recombination is independent of synapsis: implications for the immune repertoire. *Mol. Cell. Biol.* **20**:7914–7921.
77. Zhou, L., R. Mitra, P. W. Atkinson, A. B. Hickman, F. Dyda, and N. L. Craig. 2004. Transposition of hAT elements links transposable elements and V(D)J recombination. *Nature* **432**:995–1001.